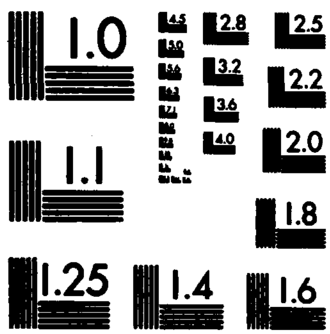
AD-A134 104 AN ALGORITHM FOR DETECTING A CHANGE IN STOCHASTIC PROCESS(U) CONNECTICUT UNIV STORRS DEPT OF ELECTRICAL ENGINEERING AND CO.. R K BANSAL ET AL. JUN 83 1/1

UNCLASSIFIED UCT/DEECS/TR-83-7 AFOSR-TR-83-0781 F/G 12/1 NL

END

FILMED



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AFOSR-TR- 83-0781

The University of Connecticut
SCHOOL OF ENGINEERING
Storrs, Connecticut 06268

AD-A134 104

(2)

An Algorithm for Detecting a Change in Stochastic Process

by

Rakesh Kumar Bansal and P. Papantoni-Kazakos
University of Connecticut
U-157
Storrs, CT. 06268

UCT/DEECS/TR-83-7

June, 1983

Department of Electrical Engineering and Computer Science

An Algorithm for Detecting a Change
in Stochastic Process

by

Rakesh Kumar Bansal and P. Papantoni-Kazakos
University of Connecticut
U-157
Storrs, CT. 06268

UCT/DEECS/TR-83-7

June, 1983

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 83-0781	2. GOVT ACCESSION NO. AD-A134 144	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) "AN ALGORITHM FOR DETECTING A CHANGE IN STOCHASTIC PROCESS"		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Rakesh Kumar Bansal; P. Papantoni-Kazakos		8. CONTRACT OR GRANT NUMBER(s) AFOSR-78-3695
9. PERFORMING ORGANIZATION NAME AND ADDRESS The University of Connecticut, School of Engineering, Storrs, Connecticut 06268		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS PE61102F; 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS AFOSR/NM, Bldg. 410, Bolling AFB DC 20332		12. REPORT DATE June 1983
		13. NUMBER OF PAGES 23
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The problem of detecting a change from one given stationary and ergodic stochastic process, to another given such process is considered. It is assumed that both the stochastic processes are processes with memory, and that they are mutually independent. A sequential test is proposed and analyzed. It is proved that the proposed test is asymptotically optimal, in a mathematically precise sense. The test utilizes a reflecting barrier at zero, and positive threshold for deciding the occurence of the change.

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

An Algorithm for Detecting a Change in
Stochastic Process

by

Rakesh Kumar Bansal and P. Papantoni-Kazakos
University of Connecticut
U-157
Storrs, Connecticut 06268

Abstract

The problem of detecting a change from one given stationary and ergodic stochastic process, to another given such process is considered. It is assumed that both the stochastic processes are processes with memory, and that they are mutually independent. A sequential test is proposed and analyzed. It is proved that the proposed test is asymptotically optimal, in a mathematically precise sense. The test utilizes a reflecting barrier at zero, and a positive threshold for deciding the occurence of the change.

This work was supported by the Air Force Office of Scientific Research under the Grant AFOSR-78-3695.

1. Introduction

Let us consider two well-known stationary processes with memory $[\mu_o, X, R]$ and $[\mu_1, Y, R]$; where μ_o and μ_1 the corresponding measures, X and Y the respective names of the processes, and R the real line on which both processes take values. Let both processes be discrete-time and let w_i^j; $j \geq i$ denote an observed data sequence $w_i, \ldots, w_j$. Let the observations start at time zero, and let it be known that the initially active process is $[\mu_o, X, R]$. Let it be possible that at some point in time, t, the process $[\mu_o, X, R]$ may become inactive, and that the process $[\mu_1, Y, R]$ may be active, instead, and remain so. Let the two processes be mutually independent, and let the time t above be allowed to take any nonnegative integer values 0,1,...,.; that is, the possibility that the process $[\mu_1, Y, R]$ becomes active at time zero is included. We consider the problem of formulating a test, that detects the $[\mu_o, X, R]$ to $[\mu_1, Y, R]$ change (if and whenever such change may occur), accurately, and fast.

The problem of detecting the possible change from a given process to another given process, has numerous applications, ranging from industrial quality control, to edge detection in images, to the diagnosis of faults in the elements of computer-communication networks. The case where both the processes $[\mu_o, X, R]$ and $[\mu_1, Y, R]$ are memoryless has been fully analyzed. Zacks and Kander [13] proposed an ad hoc procedure where fixed size n data blocks are collected, and based on those data a decision as to change is attempted. However, Page [7], [8] was the first to propose a sequential, and one-step memory algorithm, for this case. Lorden [6] proved that Page's algorithm is also asymptotically optimal, in a sense that will be explained later. In [9], [10] Page's algorithm has been successfully applied in the effective identification of faults in the links of Computer-Communication networks.

Although the problem of detecting a change in the active process has been fully analyzed in the memoryless case, no consideration has been given to the case where

the two processes $[\mu_o, X, R]$ and $[\mu_1, Y, R]$ are stationary with memory. In this paper, we undertake this task. We will first present insight to our formalization. Then, we will propose a sequential test. Finally, we will prove the asymptotic optimality of our test, in a sense that will be explained.

3. Preliminaries

Let R be the real line, and let $\mathcal{B}$ be the Borel σ-field generated by open intervals on R. Then, $(R, \mathcal{B})$ is a measurable space. Let R^n and $\mathcal{B}^n$ denote n-tuples of R and $\mathcal{B}$ respectively. Let μ^n be a probability measure defined on $(R^n, \mathcal{B}^n)$. Then, μ^n has a unique extension μ^∞ on $(R^\infty, \mathcal{B}^\infty)$. Let μ_1^∞ and μ_o^∞ be two stationary and ergodic probability measures defined on $(R^\infty, \mathcal{B}^\infty)$. Let the Kullback-Leibler information $I_{\mu_1\mu_o}(n)$ exist for all n; that is, let (1) below be true. Then, μ_1^n is aboslutely continuous with respect to μ_o^n, for all n.

$$I_{\mu_1\mu_o}(n) \triangleq \int_{R^n} \ell n \frac{d\,\mu_1^n}{d\,\mu_o^n} \cdot d\,\mu_1^n < \infty \;;\; \forall\, n. \tag{1}$$

Given probability measures μ_o^∞ and μ_1^∞ as above, let it be known that the measure μ_o^∞ is initially active. Let it be possible that at some time instant m: $m \geq 0$, the measure μ_o^∞ may become inactive, and the measure μ_1^∞ may become active instead. Thus, given an observation sequence w_o^n; $n \geq 0$, it is possible that an initial portion w_o^m of the sequence has been generated by the measure μ_o^∞, and the remaining portion w_{m+1}^n has been generated by the measure μ_1^∞; where $-1 \leq m \leq n$; if $m = -1$, the total sequence w_o^n has been generated by the measure μ_1^∞; if $m = n$, the total sequence has been generated by the measure μ_o^∞. Given some observation sequence, our objective is to decide if the μ_o^∞ to μ_1^∞ change has occurred. Since the very occurence, and even more the time, of such a change is uncertain, any decision test that we may develop should be clearly sequential. To gain insight into the problem, however, we will initially assume that a fixed length n+1 observation sequence w_o^n is available; where n: $0 \leq n \leq \infty$.

Let n be a nonnegative finite integer. Let w_o^n be an observation sequence, and let the measures μ_o^∞, μ_1^∞ be mutually independent. Given w_o^n, we want to make a decision, regarding the possibilities:

$$\left.\begin{array}{ll} w_o^m & \text{generated by } \mu_o^\infty \\ w_{m+1}^n & \text{generated by } \mu_1^\infty \end{array}\right\} \text{; For all } m : -1 \leq m \leq n \qquad (2)$$

For fixed n, the formalization (2) is a hypothesis testing problem, with n+2 hypotheses. Let us denote by H_m the hypothesis that the μ_o^∞ to μ_1^∞ change occurred just after the mth datum; where H_{-1} denotes the hypothesis that the total sequence w_o^n has been generated by the measure μ_1^∞, and where H_n denotes the hypothesis that the total sequence w_o^n has been generated by the measure μ_o^∞. Let both the measures μ_o^∞ and μ_1^∞ have density functions, and let $f_o\left(w_i \middle| w_k^j\right)$ and $f_1\left(w_i \middle| w_k^j\right)$; $j \geq k$ denote conditional density functions, induced by the measures μ_o^∞ and μ_1^∞ respectively. Then, assuming equally probable hypotheses, the optimal Bayesian formalization results in the following test:

Decide in favor of H_m if:

$$\sum_{i=m+1}^{n} \log \frac{f_1\left(w_i \middle| w_{m+1}^{i-1}\right)}{f_o\left(w_i \middle| w_o^{i-1}\right)} = \max_{-1 \leq k \leq n} \left(\sum_{i=k+1}^{n} \log \frac{f_1\left(w_i \middle| w_{k+1}^{i-1}\right)}{f_o\left(w_i \middle| w_o^{i-1}\right)} \right)$$

; where (3)

$$f_o\left(w_o \middle| w_o^{-1}\right) \triangleq f_o(w_o)$$

$$f_1\left(w_{m+1} \middle| w_{m+1}^{m}\right) \triangleq f_1(w_{m+1})$$

$$\sum_{i=n+1}^{n} \log \frac{f_1\left(w_i \middle| w_{n+1}^{i-1}\right)}{f_o\left(w_i \middle| w_o^{i-1}\right)} \triangleq 0$$

Given the density functions f_1 and f_o, let us define the following two statistics,

$$T'_n(w_o^n) \triangleq \max_{-1\leq k\leq n}\left(\sum_{i=k+1}^{n} \log \frac{f_1(w_i|w_{k+1}^{i-1})}{f_o(w_i|w_o^{i-1})}\right) \tag{4}$$

$$T_n(w_o^n) \triangleq \max_{-1\leq k\leq n}\left(\sum_{i=k+1}^{n} \log \frac{f_1(w_i|w_o^{i-1})}{f_o(w_i|w_o^{i-1})}\right) \tag{5}$$

Given an infinite data sequence w, the natural sequential extension of the test in (3) includes the statistic $T'_n(w_o^n)$ in (4), and a threshold γ, and it consists of a stopping variable $N'(w)$, defined as follows,

$$N'_\gamma(w) \triangleq \inf\{n : T'_n(w_o^n) \geq \gamma\} \tag{6}$$

Given w, the sequential test stops then at $N'_\gamma(w)$, and it is decided that the μ_o^∞ to μ_1^∞ change has occurred.

From the definitions of the statistics $T'_n(w_o^n)$ and $T_n(w_o^n)$ in (4) and (5), it is clear that an appropriate δ threshold exists, such that the stopping variable,

$$N_\delta(w) \triangleq \inf\{n : T_n(w_o^n) \geq \delta\} \tag{7}$$

maps the stopping variable $N'_\gamma(w)$ in (6). The sequential test determined by the stopping variable $N_\delta(w)$ in (7) has recursive properties that the test determined by $N'_\gamma(w)$ does not possess. Indeed, it is easily concluded from (5) that the following, recursion holds,

$$T_{n+1}(w_o^{n+1}) = \max\left(0,\ T_n(w_o^n) + g_n(w_o^{n+1})\right)$$

; where (8)

$$g_n(w_o^{n+1}) \triangleq \log \frac{f_1(w_{n+1}|w_o^n)}{f_o(w_{n+1}|w_o^n)}$$

It is clear from expressions (8) that the test determined by the stopping rule $N_\delta(w)$ is one-sided. For arbitrary stationary processes μ_o^∞, μ_1^∞, the recursions in (8) require that the whole data sequence w_o^n be kept in memory. If the processes μ_o^∞, μ_1^∞ are ℓ-order Markov, only an ℓ-size memory is required. If the processes μ_o^∞ and μ_1^∞ are memoryless, no memory is needed. In the latter case, the two stopping variables in (6) and (7) are identical (for $\gamma = \delta$), and either one is as in [7], [8], and [6]. Lorden [6] proved that then, the sequential test determined by the stopping variable $N_\delta(w)$ in (7) is asymptotically optimal; that is, for $\delta \to \infty$, we then have,

$$E_{\mu_o}\{N_\delta(w)\} \geq \delta \tag{9}$$

$$E_{\mu_1}\{N_\delta(w)\} \sim |\ell og\ \delta| \cdot I_1^{-1} \tag{10}$$

; where

$$I_1 \triangleq E_{\mu_1}\{\ell og \frac{f_1(W)}{f_o(W)}\} < \infty \tag{11}$$

and any stopping rule $N(\delta)$ that satisfies (9), is such that $E_{\mu_1}\{N(\delta)\} \geq E_{\mu_1}\{N_\delta(w)\}$.

Qualitative speaking, Lorden's result says that when the two stationary processes μ_o^∞ and μ_1^∞ are also memoryless, then, the stopping rule $N_\delta(w)$ in (7) gives for $\delta \to \infty$ the minimum possible expected time from the occurence of the μ_o^∞ to μ_1^∞ change (given that this change occurred) to its detection, among all the possible stopping rules that satisfy the false alarm bound in (9).

In this paper, we consider the case where the two processes μ_o^∞ and μ_1^∞ are stationary with memory, and ergodic. We then propose the test determined by the stopping variable $N_\delta(w)$ in (7). We will prove asymptotic optimality of the test in a precise fashion, for processes possessing certain properties. We will derive

bounds on the performance induced by one-sided tests, and we will show that those bounds are attained by the stopping rule $N_\delta(w)$ in (7), if the above properties are satisfied.

3. Performance Bounds for One-Sided Tests

In this section, we will establish performance bounds on one-sided sequential tests. To do that, we first need a theorem parallel to theorem 2 in [6].

Theorem 1

Let μ_o^∞ and μ_1^∞ be two stationary and ergodic processes. Let $w \overset{\Delta}{=} w_o, w_1, \ldots$ be an infinite data sequence, and let N be an extended stopping variable with respect to w, such that,

$$P_{\mu_o}(N < \infty) \leq \alpha \tag{12}$$

; where

$$\alpha : 0 < \alpha < 1$$

For $k = 1,2,\ldots$, let N_k denote the stopping variable obtained by applying N to $w_k, w_{k+1}, \ldots$, and define,

$$N^* = \min \{N_k + k-1 | k = 1, 2, \ldots\} \tag{13}$$

Then, N^* is an extended stopping variable, and,

$$E_{\mu_o}\{N^*\} \geq \alpha^{-1} \tag{14}$$

$$E_{\mu_1}\{N^{*^+}\} \overset{\Delta}{=} \bar{E}_{\mu_1}\{N^*\} \leq E_{\mu_1}\{N\}; \text{ for any stationary and ergodic } \mu_1^\infty. \tag{15}$$

Proof

The fact that N^* is an extended stopping variable, and expression (15) are proven exactly as in theorem 2, reference [6]. Indeed, since the events $\{N_i \leq n-i+1\}; 1 \leq i \leq n$ are determined by the sequence w_1^n, and the event $\{N^* \leq n\}$

is the union of the above, N^* is an extended stopping variable. Also, (15) holds since

$$E_{\mu_1}\{N^{*^+}|w\} \leq E_{\mu_1}\{N_m|w_1^{m-1}\} = E_{\mu_1}\{N_m\} = E_{\mu_1}\{N\}.$$

To prove expression (14), define as in [6],

$$\xi_k = \begin{cases} 1 \; ; \; N_k < \infty \\ \\ 0 \; ; \; N_k = \infty \end{cases} \; ; \; k \geq 1$$

Since μ_o^∞ is ergodic, we have,

$$\lim_{n\to\infty} n^{-1} \sum_{k=1}^{n} \xi_k = E_{\mu_o}\{\xi_1\} = P_{\mu_o}(N_1 < \infty) \leq \alpha \quad \text{a.e.}(P_{\mu_o}) \tag{16}$$

If $E_{\mu_o}\{N^*\} = \infty$, (14) is trivial. Let $E_{\mu_o}\{N^*\} < \infty$, let $N_o^* \triangleq 0$, and let N_i^*; $i \geq 1$ be such th ,

$$N_i^* < N_{i+1}^* \; ; \; \forall \, i$$

If $N_{m-1}^* = n$, apply N to $w_{n+r}, w_{n+r+1}, \ldots,$ for all r's, and define N_m^* as the first time that stopping occurs.

Then, $N_1^* = N^*$, and the variables $N_m^* - N_{m-1}^*$ are identically distributed, for different m values, no less than two. Since $\xi_{N_m^*+r} = 1$ for some r, that causes stopping at N_{m+1}^*, we have,

$$\xi_{N_m^*+1} + \ldots + \xi_{N_{m+1}^*} \geq 1$$

and

$$\xi_1 + \ldots + \xi_{N_m^*} \geq m \; ; \; m \geq 0$$

Thus,

$$\frac{\xi_1 + \ldots + \xi_{N_m^*}}{N_m^*} \geq \frac{m}{N_m^*} \; ; \; m \geq 0 \tag{17}$$

Due to (16), the left hand part of (17) is bounded from above by α, as $m \to \infty$. The right hand part of (17) converges to $E_{\mu_o}^{-1}\{N^*\}$, as $m \to \infty$, by the ergodic theorem.

We point out here, that if the extended stopping variable N is determined by the statistic $\sum_{n\geq 0} g_n(w_o^n)$; where $g_n(w_o^n)$ is given by (8), and the threshold is α^{-1}, then expression (12) in theorem 2 is satisfied [11]. Thus, theorem 1 applies to the stopping variable $N_\delta(w)$ in (7), if we put $\delta = \alpha^{-1}$. The stopping variable $N_\delta(w)$ is exactly N^* in (13), for N determined by the statistic $\sum_{n\geq C} g_n(w_o^n)$. Thus, we have,

$$E_{\mu_o}\{N_\delta(w)\} \geq \delta \tag{18}$$

$$\bar{E}_{\mu_1}\{N_\delta(w)\} \leq E_{\mu_1}\{N\} \tag{19}$$

for

$$N = \inf\left\{n \; : \sum_{i=0}^{n} g_i(w_o^i) \geq \delta\right\} \tag{20}$$

Due to theorem 1, and the conditions (18), (19), and (20), satisfied by the stopping variable $N_\delta(w)$ in (7), we will now focus our ultimate objective on showing that for $\delta \to \infty$, and under certain conditions, $\bar{E}_{\mu_1}\{N_\delta(w)\}$ is the infimum among all the expected values $\bar{E}_{\mu_1}\{N^*\}$; where N^* is any extended stopping variable that satisfies the condition $E_{\mu_o}\{N^*\} \geq \delta$. Thus, we will show that the stopping variable $N_\delta(w)$ is then optimal in the above sense. Our approach will be as follows. We will first find an upper bound on $\bar{E}_{\mu_1}\{N_\delta(w)\}$, for $\delta \to \infty$. Then we will show that this upper

bound can not be smaller asymptotically ($\delta \to \infty$) than $\bar{E}_{\mu_1}\{N^*\}$; for any extended stopping variable satisfying $E_{\mu_o}\{N^*\} \geq \delta$, and under the appropriate assumptions. Let us define,

$$L_n \triangleq n^{-1} \log \frac{f_1(w_o^{n-1})}{f_o(w_o^{n-1})} \tag{21}$$

$$I_{10} \triangleq \lim_{n\to\infty} L_n \tag{22}$$

$$p_n(\nu) \triangleq P_{\mu_1}(L_n < \nu) \tag{23}$$

Let us now impose the following conditions,

I_{10} exists ($I_{10} < \infty$) and is equal to $E_{\mu_1}\{I_{10}\}$, a.e. (P_{μ_1})

For $\nu \in (0, I_{10})$, we have,

(A)

$$\lim_{n\to\infty} np_n = 0$$

$$\sum_n p_n < \infty$$

The conditions (A) are relatively mild, and they are satisfied by most ergodic processes. If those conditions are satisfied, then we have (Berk [1]):

$$\lim_{\alpha\to 0} E_{\mu_1}\{N\} \sim \frac{|\log \alpha|}{E_{\mu_1}\{I_{10}\}}$$

or (24)

$$\lim_{\delta\to\infty} E_{\mu_1}\{N\} \sim \frac{|\log \delta|}{E_{\mu_1}\{I_{10}\}} \; ; \text{ for } \delta = \alpha^{-1}$$

; where N the extended stopping variable in (20).

We note that conditions (24) are also satisfied if instead of conditions (A), the martingale conditions of Chow and Robbins [2], or the strong mixing conditions of Lai [4] hold. From (19) and (24), we thus conclude that, for $\delta \to \infty$, and if the conditions (A) hold, the expected value $\bar{E}_{\mu_1}\{N_\delta(w)\}$ does not exceed $|\ell og\ \delta| \cdot E^{-1}_{\mu_1}\{I_{10}\}$. Therefore, $|\ell og\ \delta| \cdot E^{-1}_{\mu_1}\{I_{10}\}$ is then an upper bound on $\bar{E}_{\mu_1}\{N_\delta(w)\}$.

Let C^*_δ be the class of all the extended stopping variables N^*, satisfying $E_{\mu_o}\{N^*\} \geq \delta$. Let us define,

$$n^*(\delta) \triangleq \inf_{N^* \varepsilon C^*_\delta} \bar{E}_{\mu_1}\{N^*\} \tag{25}$$

Our final objective will be to prove that for processes μ_o^∞ and μ_1^∞ that satisfy conditions (A), we have,

$$\lim_{\delta \to \infty} n^*(\delta) \geq |\ell og\ \delta| \cdot E^{-1}_{\mu_1}\{I_{10}\} \tag{26}$$

We will undertake this task in the next section. In the remaining of this section, we will present a useful theorem and a lemma.

<u>Theorem 2</u> (Wald [11])

Let N be the sample size of a test μ_o against μ_1, with error probabilities α and β respectively. Let μ_o and μ_1 be stationary, let the conditions (A) hold, and let I_{10} be as in (22). Then,

$$E_{\mu_1}\left\{\ell og\ \frac{f_1(w_o^{N-1})}{f_o(w_o^{N-1})}\right\} \geq (1-\beta)\ \ell og\ \frac{1-\beta}{\alpha} + \beta\ \ell og\ \frac{\beta}{1-\alpha} \geq (1-\beta)\ |\ell og\ \alpha| - \ell og\ 2 \tag{27}$$

<u>Proof</u>

The first part of the above double inequality is in Wald [11].

The lower bound in the above inequality is due to $\beta \,\ell og(1-\alpha)^{-1}$ being nonnegative, and due to the fact that the minimum value of $\beta \,\ell og\, \beta + (1-\beta)\, \ell og\, (1-\beta)$ is equal to $-\ell og\, 2$.

<u>Lemma 1</u>

Let $\{a_i\}$ and $\{b_i\}$ be two sequences, such that,

$$a_i \geq 0 \;;\; \forall\, i, \sum_1^\infty a_i = 1$$

$$b_i \geq 0 \;;\; \forall\, i,\; b_i \geq b_{i+1};\; \forall\, i$$

Then,

$$\sum_{i=1}^\infty i\, a_i\, b_i \leq \left(\sum_1^\infty i\, a_i\right)\left(\sum_1^\infty a_i\, b_i\right) \tag{28}$$

<u>Proof</u>

Define,

$$c_j \triangleq \sum_j^\infty a_i\, b_i \;;\; j \geq 1$$

Then,

$$c_1 = c_1\left(\sum_1^\infty a_i\right) = c_1\, a_1 + c_1 \sum_2^\infty a_i = a_1\, b_1 + \sum_2^\infty a_i\, b_i \leq$$

$$\leq b_1 \sum_1^\infty a_i = b_1 \rightarrow a_1\, c_1 - a_1\, b_1 \leq 0 \rightarrow$$

$$\rightarrow \sum_2^\infty a_i\, b_i = c_1 \sum_2^\infty a_i + c_1\, a_1 - a_1\, b_1 \leq c_1 \sum_2^\infty a_i$$

Similarly,

$$\sum_{j}^{\infty} a_i\, b_i \le c_j \sum_{j}^{\infty} a_i \le c_1 \sum_{j}^{\infty} a_i \;;\; j \ge 2$$

Now,

$$\sum_{1}^{\infty} i\, a_i\, b_i = \sum_{k=1}^{\infty} \sum_{k}^{\infty} a_i\, b_i \le \sum_{k=1}^{\infty} c_1 \sum_{k}^{\infty} a_i = c_1 \sum_{1}^{\infty} i\, a_i$$

and the proof is complete.

4. The Infimum of the Stopping Time

In this section, we will prove that for stationary processes μ_o^∞ and μ_1^∞, that satisfy the conditions (A), the bound in (26) holds. We will do that via a theorem. In the proof of the theorem, theorem 2 and lemma 1 will be used. Note that if the limit I_{10} in (22) exists, and the processes μ_1^∞ and μ_o^∞ are also ergodic, then the limit I_{10} is the mismatch entropy of μ_1^∞ with respect to μ_o^∞.

Theorem 3

Let C_δ^* be the class of all extended stopping variables N^*, that satisfy the condition $E_{\mu_o}\{N^*\} \ge \delta$. Let $n^*(\delta)$ be as in (25), let the conditions (A) be satisfied, and let I_{10} be as in (22). Then,

$$\lim_{\delta\to\infty} n^*(\delta) \sim \frac{\log \delta}{E_{\mu_1}\{I_{10}\}}$$

Proof

It suffices to show that for every ε in (0,1), there exists $c(\varepsilon) < \infty$, such that for any stopping variable N in C_δ^*, we have,

$$E_{\mu_1}\{I_{10}\} \cdot \bar{E}_{\mu_1}\{N\} \geq (1-\varepsilon)\, \ell og\, E_{\mu_o}\{N\} - c(\varepsilon) \tag{29}$$

<u>Step 1</u>

Given ε, let us define stopping variables T_i; $i \geq 0$, such that,

$$T_o \triangleq 0\ ,\ T_i < T_{i+1} < \infty\ ;\ \forall\ i$$

T_{i+1} : the smallest n (or ∞ if such an n does not exist) such that $T_i < n$ and $f_1(w^n_{T_i+1}) \leq \varepsilon\ f_o(w^n_{T_i+1})$

From Wald's [11] argument, we have that $P_{\mu_1}(T_1 < \infty) = P_{\mu_1}(\text{decide } H_o) = \beta$. Then ε is the lowest threshold for deciding in favor of H_o, which in Wald's test is larger than $\beta(1-\alpha)^{-1} > \beta$. Thus, $P_{\mu_1}(T_1 < \infty) \leq \varepsilon$. Let us define,

$$D_{rk} \triangleq \{w^k_1 : T_{r-1} = k\}\ ;\ k < N \tag{30}$$

Let us denote by P_{k+1} the probability induced by the measure μ^∞_1 and applied to data sequences $w_{k+1}, w_{k+2}, \ldots$. Then, provided that $P_1(D_{rk}) > 0$, we have due to the arguments above: $P_{k+1}(T_r < \infty | D_{rk}) \leq \varepsilon$. On the set D_{rk} in (30), and for given N in C^*_δ, given T_r, we define a sequential test based on the sequence $w_{k+1}, w_{k+2}, \ldots,$ as follows:

$$\text{Stop at min } (N, T_r) \cdot$$

$$\text{Decide} : \begin{cases} P_{k+1} & ; \text{ if } N \leq T_r \\ P_{\mu_o} & ; \text{ if } N > T_r \end{cases}$$

The number of observations taken for the above test is min $(N, T_r) - k$,

whose conditional expectation $P_{k+1}(\cdot|D_{rk})$ is at most $\bar{E}_{\mu_1}\{N\}$. To this point, our derivations are basically as in the proof of theorem 3 in Lorden [6], with the appropriate modifications.

Step 2

Now, Wald' inequality (27) in theorem 2 holds, on D_{rk} with $\alpha = P_{\mu_o}(N \le T_r|D_{rk})$ and $1-\beta = P_{k+1}(N \le T_r|D_{rk})$. So, we have,

$$E_{k+1}\left\{\ell og\, \frac{f_1(w_{k+1}^{N'})}{f_o(w_{k+1}^{N'})}\middle| D_{rk}\right\} \ge P_{k+1}(N \le T_r|D_{rk})\cdot|\ell og\, P_{\mu_o}(N \le T_r|D_{rk})| - \ell og\, 2 \ge$$

$$\ge (1-\varepsilon)|\ell og\, P_{\mu_o}(N \le T_r|D_{rk})| - \ell og\, 2 \tag{31}$$

; where $N' \triangleq \min(N, T_r)$ and the last part in (31) is due to,

$P_{k+1}(N \le T_r|D_{rk}) \ge P_{k+1}(T_r = \infty|D_{rk}) \ge 1-\varepsilon$

Due to the conditions (A), given $\xi > 0, \exists\, N_o(\xi) < \infty$:

$$\left|E_{k+1}\left\{(i-k)^{-1}\, \ell og\, \frac{f_1(w_{k+1}^{i})}{f_o(w_{k+1}^{i})}\right\} - E_{\mu_1}\{I_{10}\}\right| < \xi\ ;\ \forall\, i > k + N_o(\xi) \tag{32}$$

Due to (32), we now obtain,

$$E_{k+1}\left\{\ell og\, \frac{f_1(w_{k+1}^{N'})}{f_o(w_{k+1}^{N'})}\middle| D_{rk}\right\} \le$$

$$\le \sum_{i=k+1}^{k+N_o(\xi)} P_{k+1}(N = i|D_{rk})\, E_{k+1}\left\{\ell og\, \frac{f_1(w_{k+1}^{i})}{f_o(w_{k+1}^{i})}\middle| N = i,\, D_{rk}\right\}$$

$$+ [\xi + E_{\mu_1}\{I_{10}\}]\, \sup_k\ \text{ess sup} \sum_{i=k+N_o(\xi)+1}^{\infty} (i-k)\, P_{k+1}\{N = i|w_1^k\} \le$$

$$\leq \bar{E}_{\mu_1}\{N\} \cdot [\xi + E_{\mu_1}\{I_{10}\}] + \sum_{i=k+1}^{k+N_o(\xi)} P_{k+1}(N = i|D_{rk})\, E_{k+1}\left\{\ell og \frac{f_1(w_{k+1}^i)}{f_o(w_{k+1}^i)}\middle| N = i,\, D_{rk}\right\}$$

(33)

But, due to conditions (A), the expected values $E_{\mu_1}\left\{n^{-1}\,\ell og \frac{f_1(w_1^n)}{f_o(w_1^n)}\right\}$ are bounded for all n. If this bound is B, then $E_{\mu_1}\left\{\ell og \frac{f_1(w_1^n)}{f_o(w_1^n)}\right\} < n\,B < \infty$; for $n < \infty$. Therefore, $E_{\mu_1}\left\{\ell og \frac{f_1(w_{k+1}^{k+n})}{f_o(w_{k+1}^{k+n})}\middle| w_1^k\right\}$ is bounded a.e. (P_{μ_1}).

Due to the above arguments, and since $N_o(\xi)$ is finite, we conclude that there exists some finite constant $C(\xi)$ such that,

$$\sum_{i=k+1}^{k+N_o(\xi)} P_{k+1}(N=i|D_{rk})\, E_{k+1}\left\{\ell og \frac{f_1(w_{k+1}^i)}{f_o(w_{k+1}^i)}\middle| N = i,\, D_{rk}\right\} < C(\xi)$$

; for almost all w_1^k in measure μ_1. (34)

From (33), and (34), we obtain,

$$E_{k+1}\left\{\ell og \frac{f_1(w_{k+1}^{N'})}{f_o(w_{k+1}^{N'})}\middle| D_{rk}\right\} \leq \bar{E}_{\mu_1}\{N\} \cdot [\xi + E_{\mu_1}\{I_{10}\}] + C(\xi)$$

; a. e. in D_{rk}, in μ_1

$$\underset{\xi\to 0}{\to} \bar{E}_{\mu_1}\{N\} \cdot E_{\mu_1}\{I_{10}\} + C'(\xi) \qquad (35)$$

<u>Step 3</u>

Let R be the smallest integer $r \geq 1$ such that $T_r > N$; where $\{T_i\}$ the sequence in step 1. If $P_{\mu_o}(R \geq r) > 0$, then $P_{\mu_o}(R < r + 1|R \geq r)$ is well defined and it equals $P_{\mu_o}(N \leq T_r|T_{r-1} < N)$, which is an average over k of the probabilities

$P_{\mu_o}(N \leq T_r | T_{r-1} = k < N)$, satisfying (31). Therefore, $P_{\mu_o}(R \geq r) > 0$, and convexity of $-\ell og$ implies,

$$\sum_{k=0}^{\infty} P_{k+1}(D_{rk}) \cdot E_{k+1}\left\{\ell og \frac{f_1(w_{k+1}^{N'})}{f_o(w_{k+1}^{N'})}\middle| D_{rk}\right\} \geq (1-\varepsilon)\,|\ell og\ P_{\mu_o}(R < r+1 | R \geq r)| - \ell og\ 2 \tag{36}$$

If $P_{\mu_o}(R < r+1 | R \geq r) \geq Q$; $r \geq 1$, such that $P_{\mu_o}(R \geq r) > 0$, then $P_{\mu_o}(R \geq r+1) \leq (1-Q)^r$; hence $E_{\mu_o}\{R\} \leq Q^{-1}$. Thus, we obtain from (36):

$$\sum_{k=0}^{\infty} P_{k+1}(D_{rk})\ E_{k+1}\left\{\ell og \frac{f_1(w_{k+1}^{N'})}{f_o(w_{k+1}^{N'})}\middle| D_{rk}\right\} \geq (1-\varepsilon)\ E_{\mu_o}\{R\} - \ell og\ 2 \tag{37}$$

We observe that due to (35) and (37), $E_{\mu_o}\{R\}$ is bounded, that is $E_{\mu_o}\{R\} < \infty$. Also, from (35) and (37), we obtain,

$$(1-\varepsilon)\ E_{\mu_o}\{R\} - \ell og\ 2 \leq \bar{E}_{\mu_1}\{N\} \cdot [\xi + E_{\mu_1}\{I_{10}\}] + C(\xi) \to$$

$$\underset{\xi \to 0}{\to} \bar{E}_{\mu_1}\{N\} \cdot E_{\mu_1}\{I_{10}\} + C'(\xi) \tag{38}$$

Step 4

The sequence $\{T_i\}$ in step 1 is a sequence of cumulative sums of ergodic and stationary integer valued random variables, under the probability measure μ_o. We have,

$$E_{\mu_o}\{T_R\} = \sum_{i=1}^{\infty} P_{\mu_o}(R = i)\ E_{\mu_o}\{T_R | R = i\}$$

$$= \sum_{i=1}^{\infty} P_{\mu_o}(R = i) \cdot E_{\mu_o}\{T_1 | R = i\} \cdot i \tag{39}$$

But the expected value $E_{\mu_o}\{T_1|R = i\}$ decreases with increasing i and lemma 1 applies. Thus, we obtain from (39),

$$E_{\mu_o}\{T_r\} \le \left[\sum_{i=1}^{\infty} i\, P_{\mu_o}(R = i)\right]\left[\sum_{i=1}^{\infty} P_{\mu_o}(R = i)\, E_{\mu_o}\{T_1|R = i\}\right]$$

$$= E_{\mu_o}\{R\} \cdot E_{\mu_o}\{T_1\} \tag{40}$$

; with equality if the processes μ_o, μ_1 are memoryless, which is Page's and Lorden's case.

From (40), and due to the definition of R in step 3, we obtain,

$$\ell og\ E_{\mu_o}\{N\} \le \ell og\ E_{\mu_o}\{T_R\} \le \ell og\ E_{\mu_o}\{R\} + \ell og\ E_{\mu_o}\{T_1\}$$

or

$$\ell og\ E_{\mu_o}\{R\} \ge \ell og\ E_{\mu_o}\{N\} - \ell og\ E_{\mu_o}\{T_1\} \tag{41}$$

From (38) and (41) we now obtain directly,

$$(1-\varepsilon)\ \ell og\ E_{\mu_o}\{N\} - (1-\varepsilon)\ \ell og\ E_{\mu_o}\{T_1\} - \ell og\ 2 \le$$

$$\le \bar{E}_{\mu_1}\{N\}\cdot[\xi + E_{\mu_1}\{I_{10}\}] + C(\xi)$$

$$\underset{\xi \to 0}{\to} \bar{E}_{\mu_1}\{N\}\cdot E_{\mu_1}\{I_{10}\} + C'(\xi)$$

or

$$\bar{E}_{\mu_1}\{N\}\cdot E_{\mu_1}\{I_{10}\} \ge (1-\varepsilon)\ \ell og\ E_{\mu_o}\{N\} - \left[(1-\varepsilon)\ \ell og\ E_{\mu_o}\{T_1\} + \ell og\ 2 + C'(\xi)\right] \tag{42}$$

But $E_{\mu_o}\{T_1\}$ is finite, and it does not depend on N, but only on ε in step 1. Also $C'(\xi)$ is finite and independent of ε. We can thus write

$$c(\varepsilon) \triangleq (1-\varepsilon)\ \ell og\ E_{\mu_o}\{T_1\} + \ell og\ 2 + C'(\xi) < \infty$$

and we have proven inequality (29).

The proof of the theorem is now complete.

---

5. Concluding Remarks

Via the derivations in section 3, and theorem 3, we have basically proved the following theorem.

Theorem 4

Let μ_o and μ_1 be two stationary, ergodic, and mutually independent stochastic processes with memory. Let conditions (A) in section 3 be satisfied. Consider the extended stopping variables N^*, generated by testing a μ_o to μ_1 change. Let C^*_δ be the class of all such extended stopping variables N^*, that also satisfy the condition $E_{\mu_o}\{N^*\} \geq \delta$. Let $n^*(\delta)$ be defined as follows,

$$n^*(\delta) \triangleq \inf_{N^* \varepsilon C^*_\delta} \bar{E}_{\mu_1}\{N^*\}$$

Let I_{10} be defined as follows,

$$I_{10} \triangleq \ell im_{n\to\infty}\ n^{-1}\ \ell og\ \frac{f_1(w_o^{n-1})}{f_o(w_o^{n-1})}$$

Then, due to conditions (A), I_{10} exists a.e. (P_{μ_1}) and is equal to $E_{\mu_1}\{I_{10}\}$. Furthermore,

$$\ell im_{\delta\to\infty}\ n^*(\delta) \sim \frac{\ell og\ \delta}{E_{\mu_1}\{I_{10}\}}$$

and for the extended stopping variable $N_\delta(w)$ in (7):

$$\lim_{\delta\to\infty} \bar{E}_{\mu_1}\{N_\delta(w)\} \sim \frac{\log \delta}{E_{\mu_1}\{I_{10}\}}$$

What theorem 4 basically says is that, the sequential test described by the stopping variable $N_\delta(w)$ in (7), is asymptotically ($\delta\to\infty$) optimal among all tests in class C^*_δ in the theorem. The test then minimizes the expected time between the occurence of a μ_o to μ_1 change and its detection, under the constraint that if this change does not occur, then the expected time for a false alarm (exceeding the upper threshold δ) is no less than the threshold value δ.

The sequential test described by the stopping variable $N_\delta(w)$ in (7) operates in a way exhibited by expressions (8). The statistic $T_n(w_o^n)$ is updated sequentially as follows.

$$T_{n+1}(w_o^{n+1}) = \max\left(0,\ T_n(w_o^n) + \log \frac{f_1(w_{n+1}|w_o^n)}{f_o(w_{n+1}|w_o^n)}\right)$$

The updating step is $\log \dfrac{f_1(w_{n+1}|w_o^n)}{f_o(w_{n+1}|w_o^n)}$, and the μ_o^∞ to μ_1^∞ change is decided the first time N that the statistic $T_N(w_o^N)$ exceeds the threshold δ.

We note that the conditions (A) for optimality, hold for a large class of ergodic and stationary processes. As an example, let μ_1^∞ and μ_o^∞ be both Gaussian with common spectral density and means equal to Θ and 0 respectively. Let R_n denote the n-dimensional covariance matrix induced by the common spectral density, and let R_n^{-1} be its inverse. Let $f^{-1}(\omega)$ denote the spectral density induced by R_n^{-1}, for $n\to\infty$. Then, if $f^{-1}(\omega)$ and $(2\pi)^{-1} \int_{-\pi}^{\pi} f^{-1}(\omega) \sin\omega\, [1-\cos\omega]^{-1}\, d\omega$ exist, the conditions (A) are satisfied. Furthermore, the updating step $\log \dfrac{f_1(w_{n+1}|w_o^n)}{f_o(w_{n+1}|w_o^n)}$ has then a linear from $\Theta\cdot\rho_n[w_{n+1} - h_n(w_o^n)]$. The constant ρ_n and the function $h_n(w_o^n)$ can be updated themselves sequentially, if the spectral density function of the processes μ_1^∞ and μ_o^∞ has convenient form.

References

[1] R.H. Berk (1973), "Some Asymptotic Aspects of Sequential Analysis," Ann. Statist. Vol. 1, No. 6, 1126-1136.

[2] Y.S. Chow and H. Robbins (1963), "A Renewal Theorem for Random Variables which are Dependent or Non-Identically Distributed," Ann. Math. Statist., Vol. 34, 390-395.

[3] A.N. Kolmogorov and Yu. A. Rozanov (1960), "On Strong Mixing Conditions for Stationary Gaussian Processes," Theory of Probability and Applications, Vol. V, No. 2, 204-208.

[4] Tze Leung Lai (1977), "Convergence Rates and r-Quick Versions of the Storng Law for Stationary Mixing Sequences," Ann. Prob., Vol. 5, No. 5, 693-706.

[5] M. Loéve (1963), Probability Theory, Van Nostrand, Princeton.

[6] G. Lorden (1971), "Procedures for Reacting to a Change in Distribution," Ann. Math. Statist., Vol. 42, No. 6, 1897-1908.

[7] E.S. Page (1954), "Continuous Inspection Schemes," Biometrika, Vol. 41, 100-115.

[8] E.S. Page (1955), "A Test for a Change in a Parameter Occurring at an Unknown Point," Biometrika, Vol. 42, 523-527.

[9] P. Papantoni-Kazakos (1979), "Algorithms for Monitoring Changes in Quality of Communication Links," IEEE Trans. Comm., Vol. COM-27, No. 4, 682-692.

[10] P. Papantoni-Kazakos (1979), "The Potential of End-to-End Observations in Trouble Localization and Quality Control of Network Links," IEEE Trans. Comm., Vol. COM-27, No. 1, 16-30.

[11] A. Wald (1947), Sequential Analysis, Wiley, New York.

[12] S.P. Wong (1968), "Asymptotically Optimum Properties of Certain Sequential Tests," Ann. Math. Statist., Vol. 39, 1244-1263.

[13] S. Zacks and Z. Kander (1966), "Test Procedures for Possible Changes in Parameters of Statistical Distributions Occurring at Unknown Time Points," Ann. Math. Statist., Vol. 37, 1196-1210.

END

FILMED

11-83

DTIC